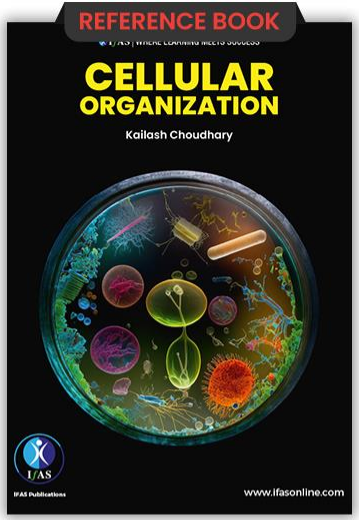
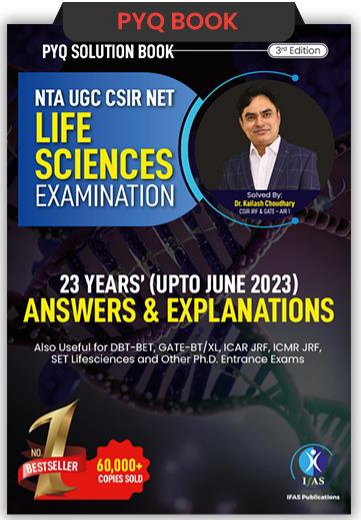
Don't wait for the opportunity. Create it.

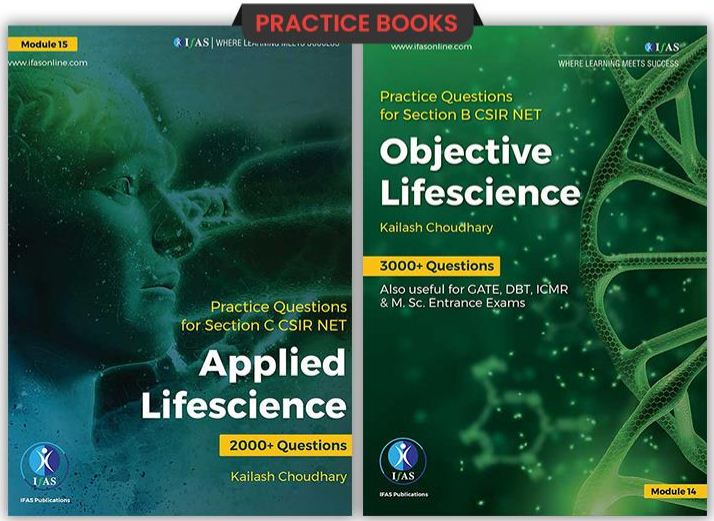
CSIR NET – Life Science

Unit 2: Cellular Organization

L11 **Genome Organization**




PYQ BOOK
PYQ SOLUTION BOOK
3rd Edition
NTA UGC CSIR NET
LIFE SCIENCES
EXAMINATION
23 YEARS' (UPTO JUNE 2023)
ANSWERS & EXPLANATIONS
Also Useful for DBT-BET, GATE-BT/XL, ICAR JRF, ICMR JRF, SET Lifesciences and Other Ph.D. Entrance Exams
NO 1 BESTSELLER
60,000+ COPIES SOLD



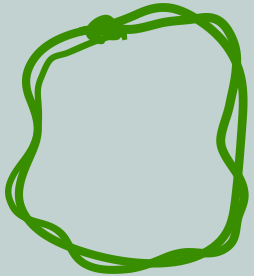
IFAS

Order Online and Get Free Delivery Across India

Points to be covered in this Lecture

- Prokaryotic Genome organization
- Eukaryotes-nucleosome, 30 nm, chromatid
- Nature of genome: c-value paradox, unique and repetitive DNA

Prokaryotic Genome organization



Bacterial genome is double stranded closed circular DNA

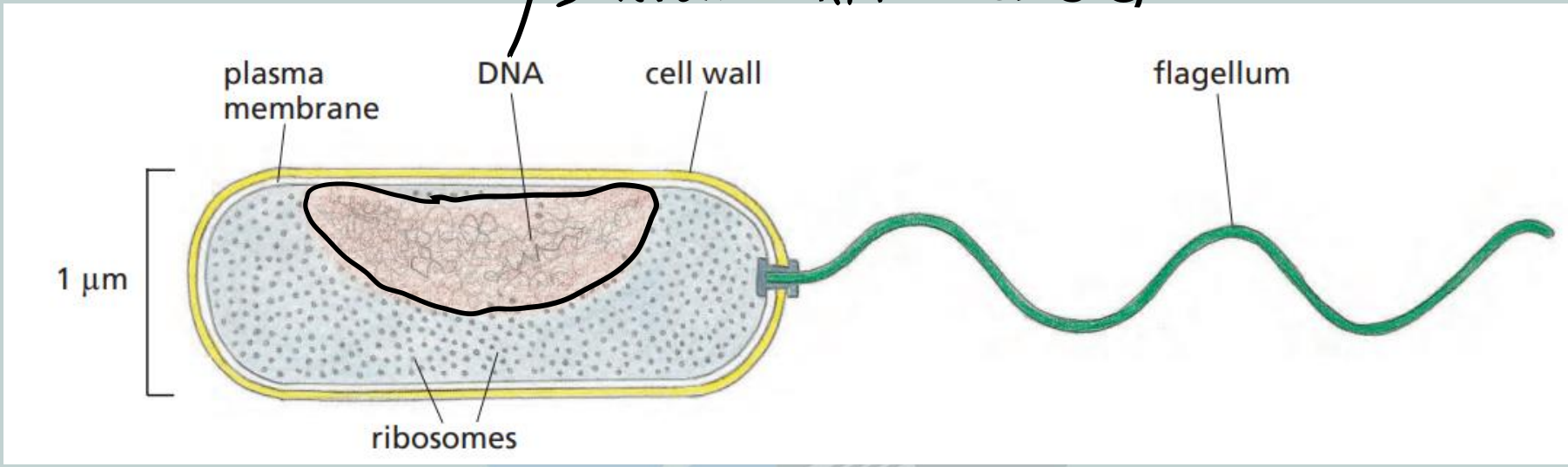
Eg. *E.coli*

Genome size = 4.2×10^6

Almost all genes are expressible except – ori c (participate in relication)

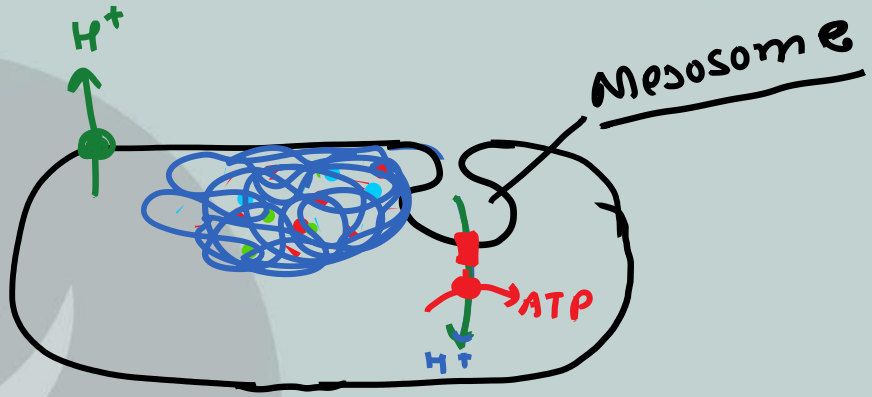
Prokaryotic gene – clustered gene (interon absent)
- polycistronic (single pramoter for multiple gene)
eg. Lac operon

Genome = Kla Nucleoid



Nucleoid

- Contain heat nuclear stable protein
- Filament temp. sensitive [fts]
- Heat nuclear unstable
- Gyrase (topo II)
  - Histone Absent



mesosome

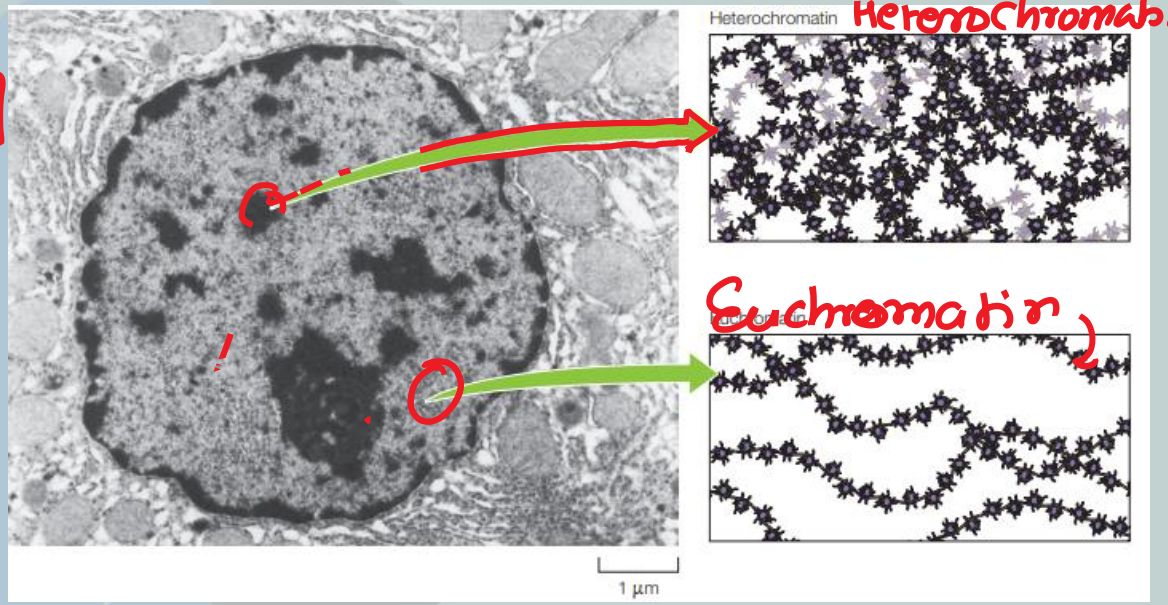
- Contain Fo -F1 particle
- Generate ATP
- Helps in replication

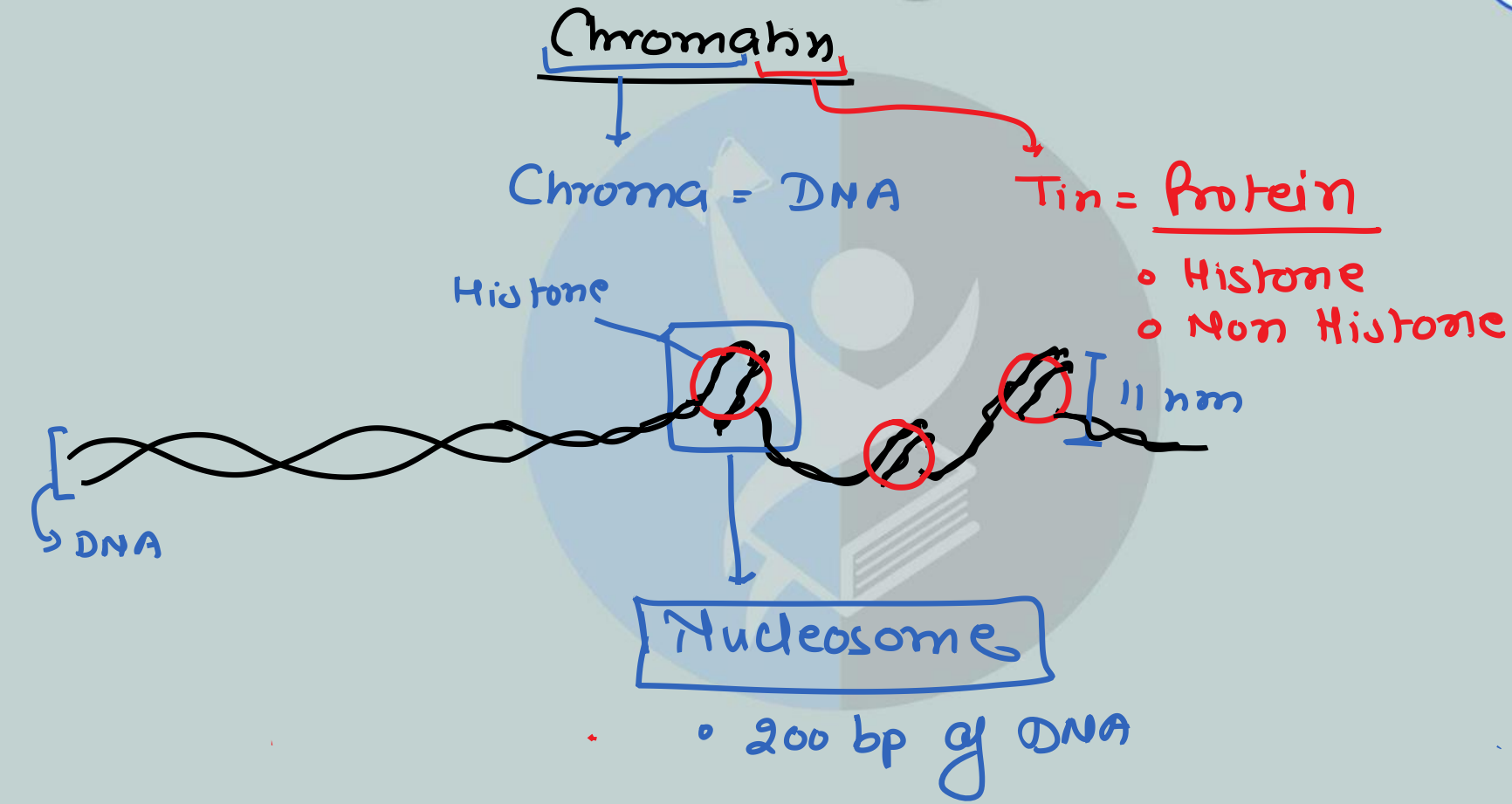
Eukaryotic Genome organization → Inter phase Nucleus

Euchromatin:
Active region [Transcription ↑]
- Stained with basic dye hematoxylin
- Less stained, light region

- Heterochromatin:

- Transcriptionally inactive
- Highly condensed
- Dark stained region



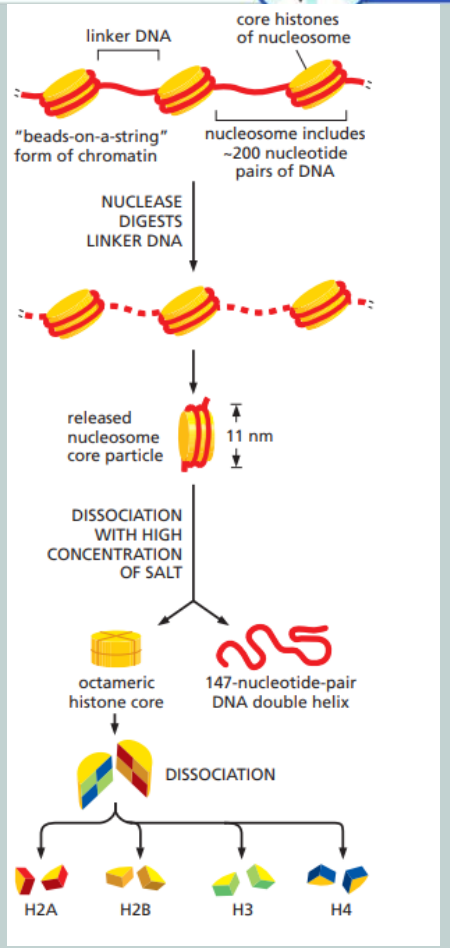

Chromatin
Chroma = DNA
Tin = Protein
o Histone
o Non Histone
Histone
11 nm
DNA
Nucleosome
o 200 bp of DNA

Structural organization of the nucleosome.

- A nucleosome contains a protein core made of eight histone molecules.
- The nuclease can degrade the exposed linker DNA but cannot attack the DNA wound tightly around the nucleosome core.
- After dissociation of the isolated nucleosome into its protein core and DNA, the length of the DNA that was wound around the core can be determined.
- This length of 147 nucleotide pairs is sufficient to wrap 1.7 times around the histone core.

53 bp = Linker DNA

1.64 Turn

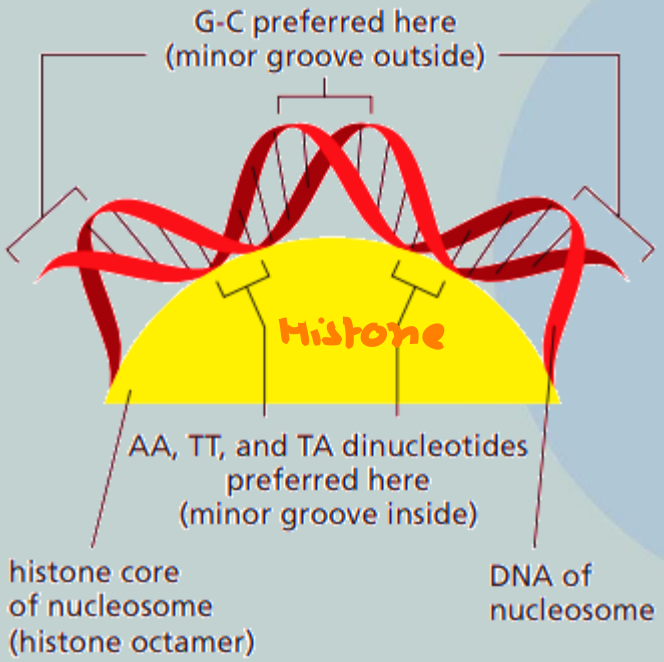


Histone Composition

- Histones are rich in lysine and arginine
- In prokaryotic genomes, histones are absent
- In sperm: - Instead of histone, protamine is present
- Protamine is rich in arginine and is a basic protein

↳ more basic than Normal Histone

The binding of DNA in a nucleosome



Interaction between Histone and DNA

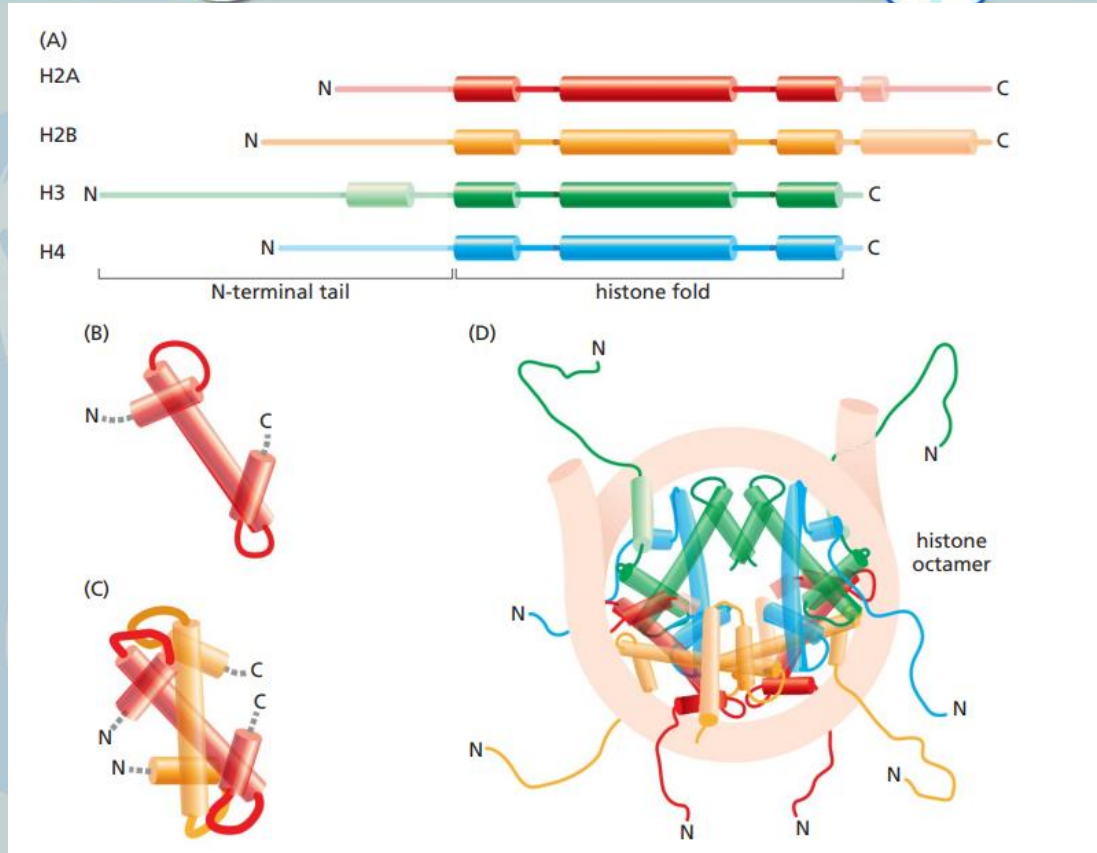
- Electrostatic interaction between histone and DNA
- AT Rich sequence DNA interacts with histone from its minor groove

(A) Each of the core histones contains an N-terminal tail, which is subject to several forms of covalent modification, and a histone fold region, as indicated.

(B) The structure of the histone fold, which is formed by all four of the core histones.

(C) Histones 2A and 2B form a dimer through an interaction known as the "handshake." Histones H3 and H4 form a dimer through the same type of interaction.

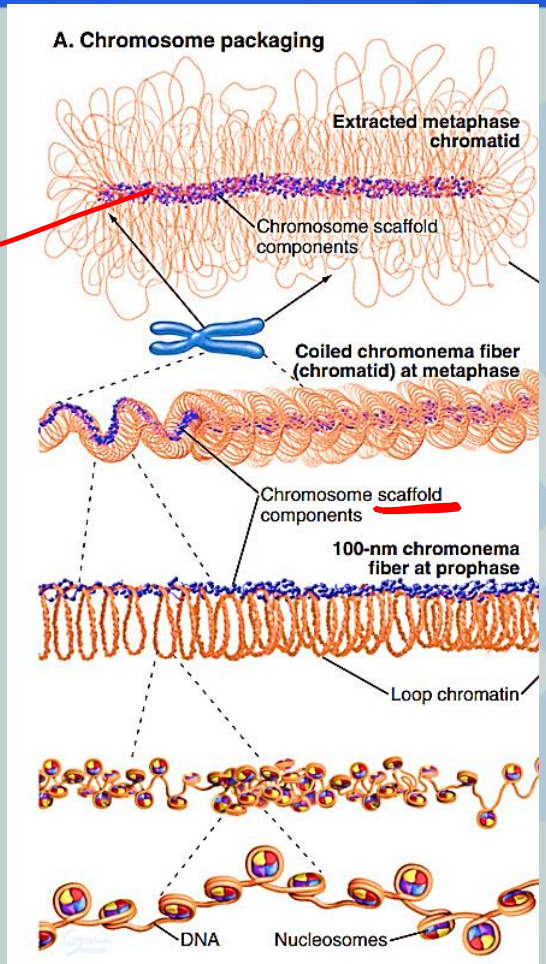
(D) The final histone octamer on DNA.

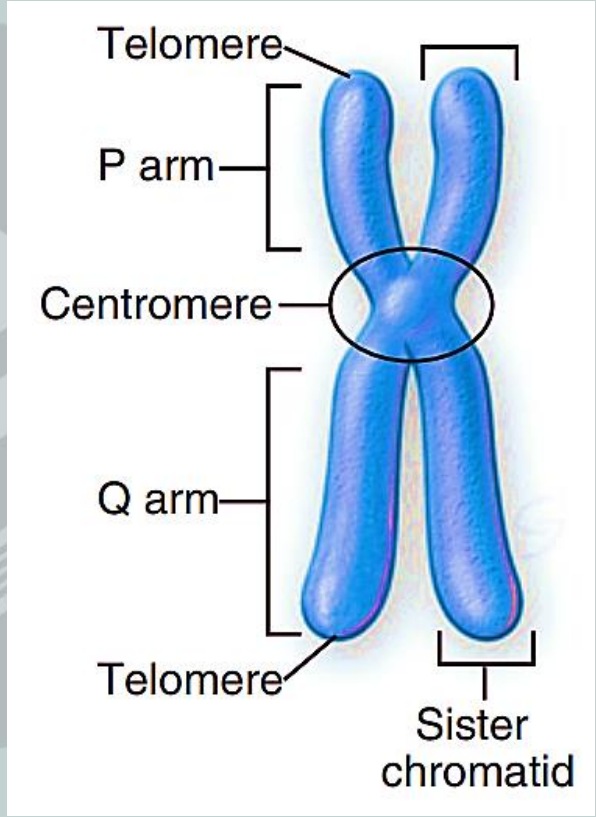


Note that all eight N-terminal tails of the histones protrude from the discshaped core structure.

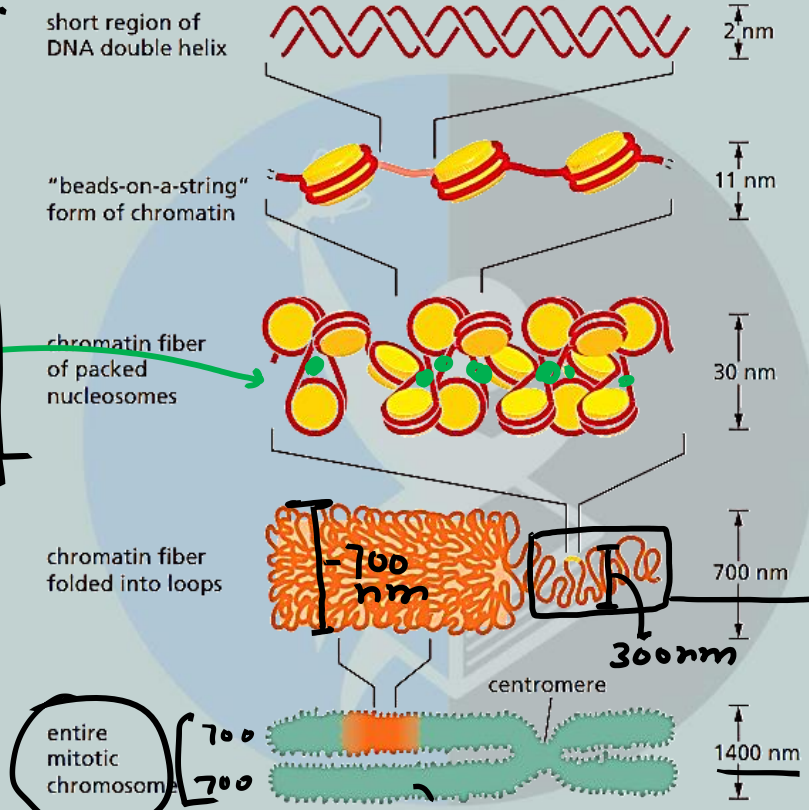
Their conformations are highly flexible, and they serve as binding sites for sets of other proteins

Scaffold Protein





Interphase



Nucleosome - Euchromatin

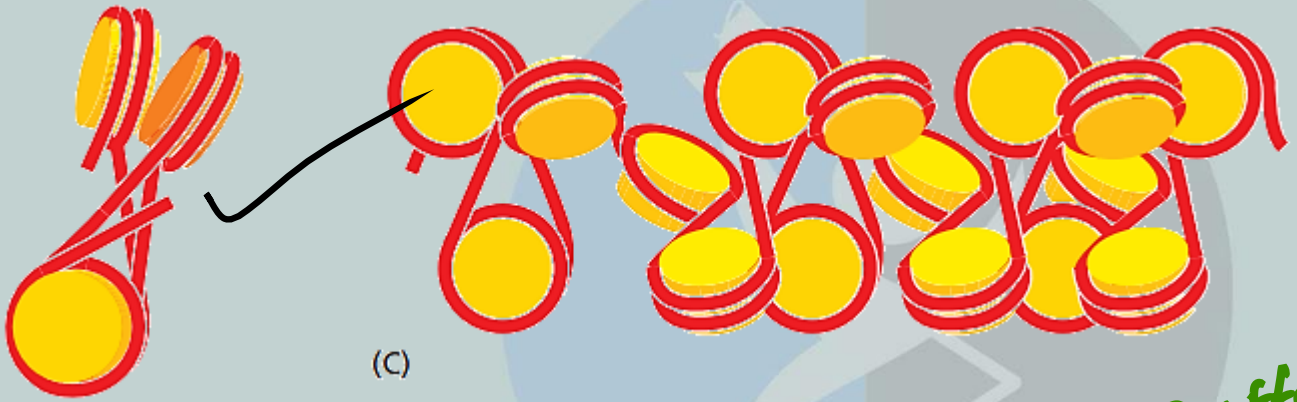
Chromatosome - Hetero Chromatin

Linker Histone (H1 Histone)
↓
Formation of Chromatosome or 30 nm fiber

looped domain → coiled of 300 nm around Non Histon
↓
form 700nm loop
↓
Chromatid

m-phase

A zigzag model for the 30- nm chromatin fiber



(C)

Zig-zag Arrangement

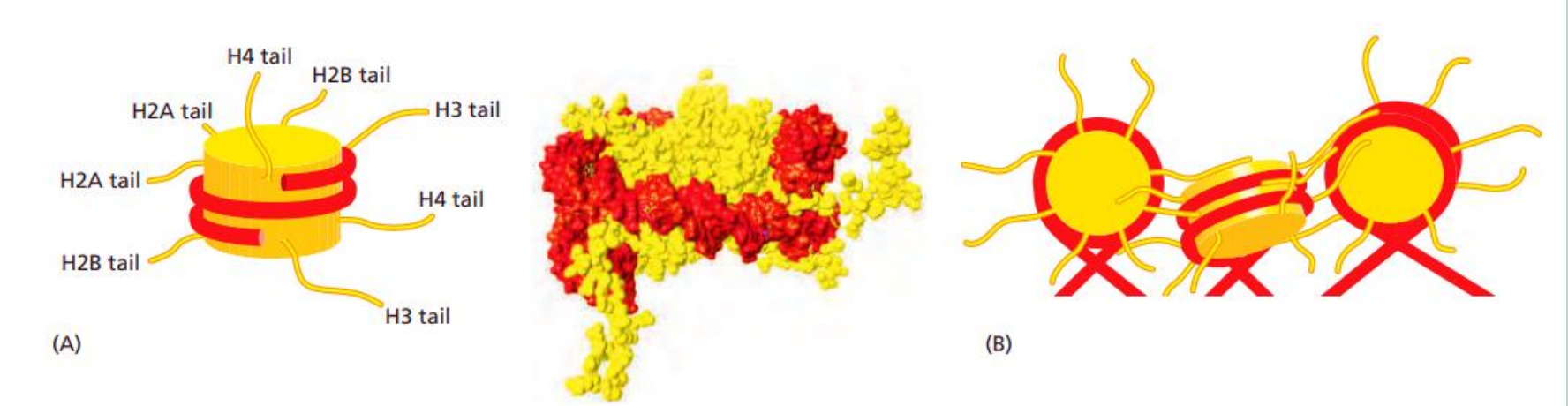
Scaffold protein

looped domain

Solenoid model

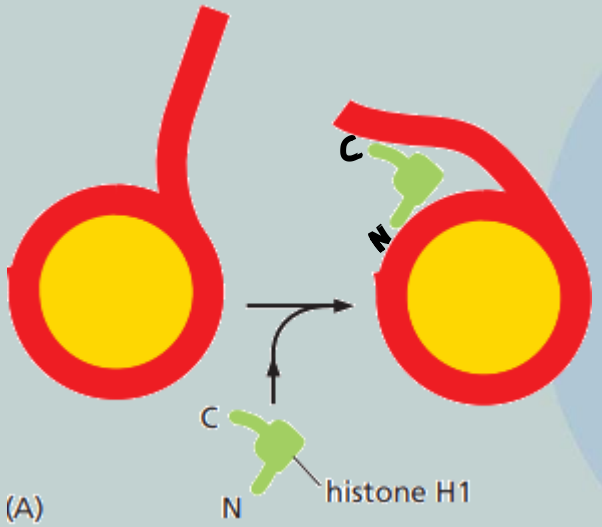
30nm fiber → Failed ✗

6 Nucleosome = Parallel Arrangement



As indicated, the histone tails are thought to be involved in interactions between nucleosomes that help to pack them together.

How the linker histone binds to the nucleosome



linker histone, H1 histone.

This is a variable histone,

H5 histone present in bird instead of H1.

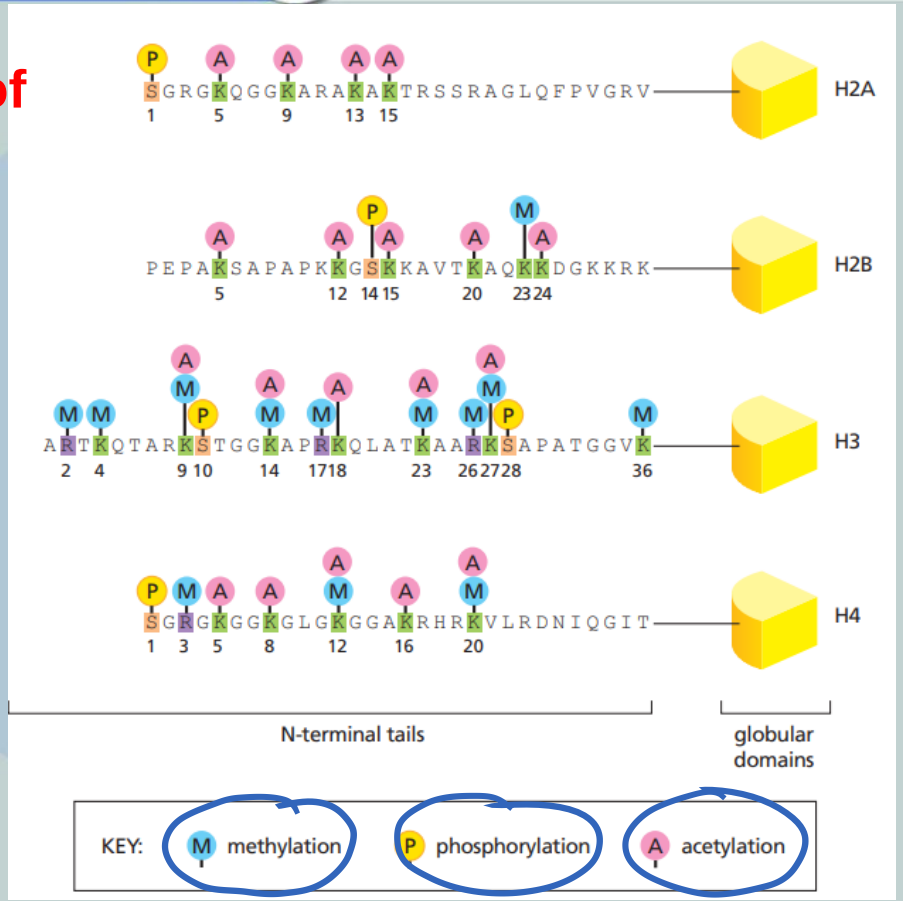
The covalent modification of core histone tails

H_2A = (P) → Ser
(A) → K

H_2B = (P) - S , (A) - K
(M) - K

H_3 = (M) → K/R
(A) → K
(P) → S

H_4 → (P) - S , (A) - K
(M) → K/R



In euchromatin, N-terminus is outside and modification can be seen on N-terminus acetylation.

↓

Acetyl group added on lysine residue of histone N-terminus.

↓

Due to acetylation, positive charge effect of histone is reduced,

↓

loose interaction between histone and DNA, euchromatin formation,

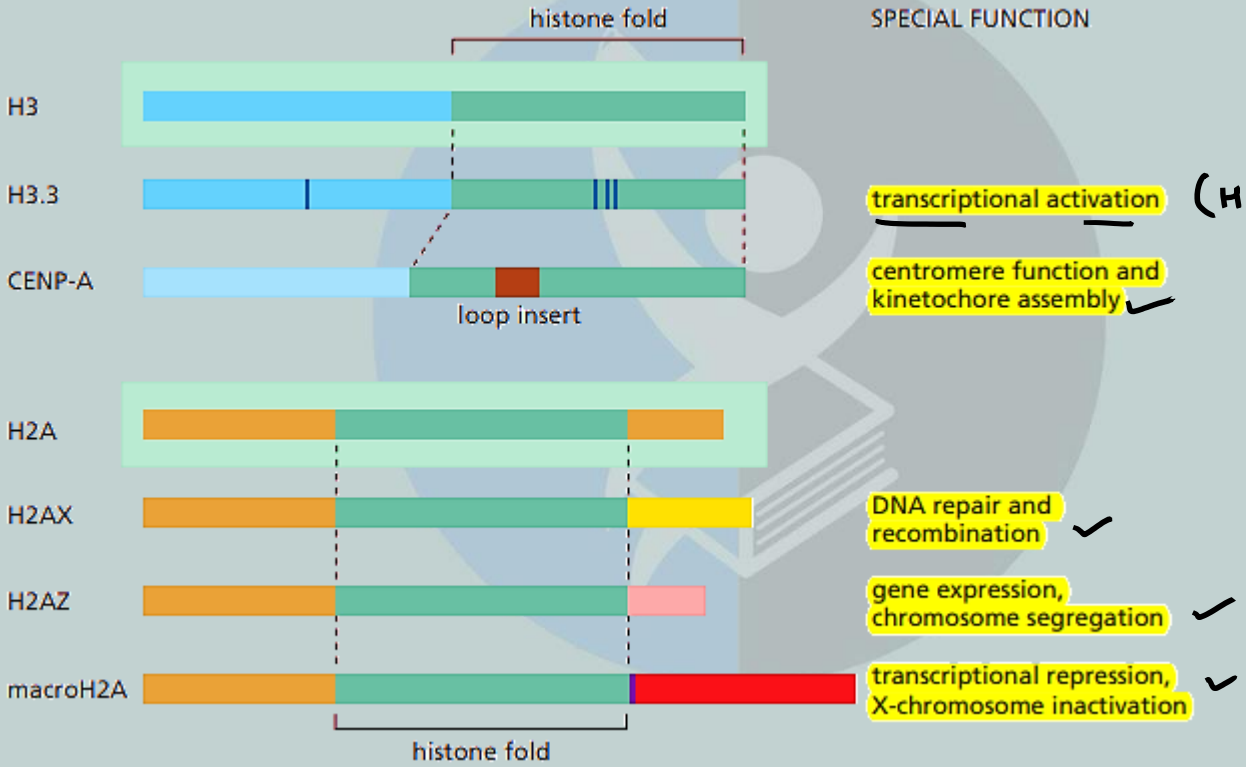
Generally - Acetylation

Hetero Chromatin → Euchromatin

Methylation

Euchromatin → Hetero Chromatin

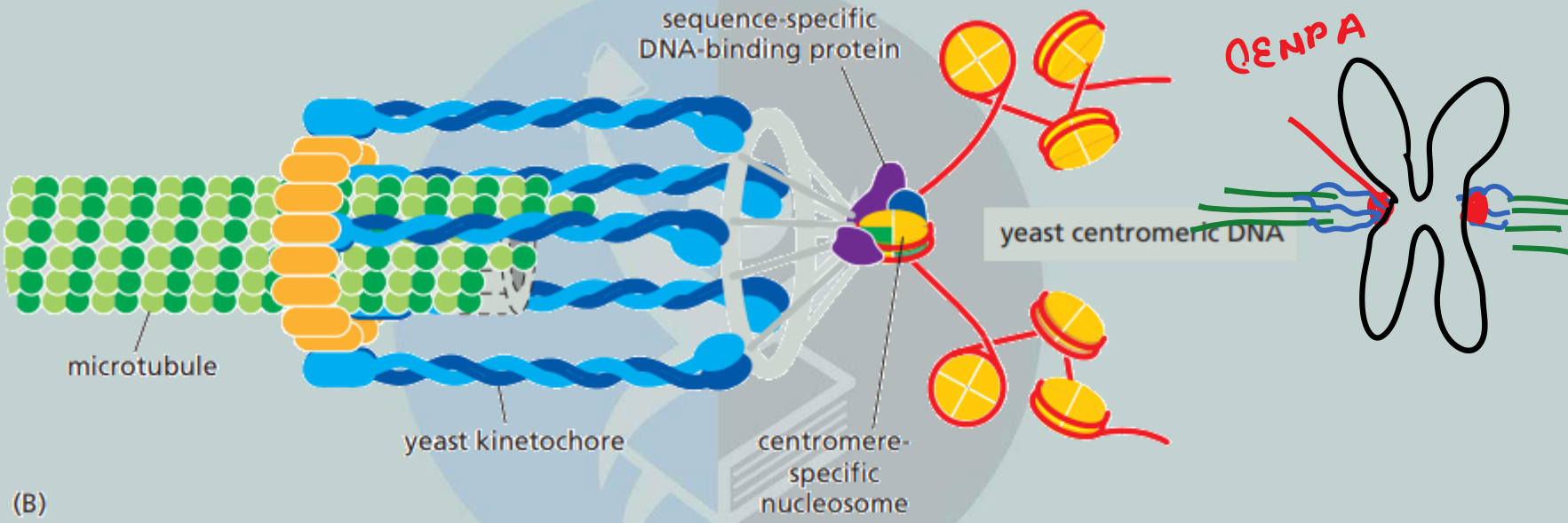
The structure of some histone variants compared with the major histone that they replace



Hu Histone is conserved

(H3.3)

- **The CENP-A (Centromere Protein-A) variant of histone H3.**



chromatin contains a centromere-specifc variant H3 histone, known as CENP-A (Centromere Protein-A; plus additional proteins that pack the nucleosomes into particularly dense arrangements and form the kinetochore, the special structure required for attachment of the mitotic spindle

Nucleosome arranged in zigzag manner,

looped domain structure,

scaffold protein - non-histone protein,

- matrix-associated region MAR,
- scaffold-associated region SAR,
- structural maintenance of chromatin SMC, e.g. Topo II

] - Helps to maintain Str. of Chromosome.

. **Euchromatin**,

- 11nm diameter,

↓

Transcriptionally Active Region

heterochromatin,

- chromatosome,
- H1 histone,
- 30nm fiber,
- 300nm looped domain structure,
- 700nm chromated,
- 1400nm chromosome.

– Inactive

Chromatid or Chromosome → Appear in M - phase

Nature of genome,

- **genome** is sum total of haploid set of chromosome,
- amount of genome is known as C-Value,
- genome size = C-Value,
- during evolution genome size or increase in C-Value.
- **C-Value paradox**, number of gene are not increase as that much as increase in genome size,
- reason extragenic DNA or Junk DNA.

During Evolution -

- Genome size increases = gene density decreases.

Organism	Genome size (Mb)[a]	Number of genes
Bacteria		
E. coli	4.6	4288
Yeasts		
S. cerevisiae	12	6000
Invertebrates		
C. elegans	97	19,000
Drosophila	180	13,600
Plants		
Arabidopsis thaliana	125	26,000
Mammals		
Human	3200	21,000

[a]Mb = millions of base pairs
[b]NS = not significant

Human genome size = 3 x 10⁹ base pair

Out of which 40% is genic or gene related sequence

- protein coding reason is only - 1 to 1.5%,
- Rest is non-coding reason e.g. 5'-UTR, 3'-UTR, promoter, pseudogene.

tRNA, rRNA, miRNA
Regulatory RNA

60% extragenic sequence out of which

- 10% is unique sequence and
- 50% is repetitive sequence- ,
- 45% large sized genome wide repeats
- e.g. LINE and SINE,
- 5% tandem repeats e.g. satellite DNA

Pseudogene and its ancestral connection

- A **pseudogene** is a DNA sequence closely resembling that of a functional gene, but containing numerous mutations that prevent its proper expression or function.

- Most pseudogenes arise from the duplication of a functional gene followed by the accumulation of damaging mutations in one copy.

. If there is reversal of mutation occurs in pseudogene (Ancestral gene)

↓

pseudogene expression ON.

↓

Reversal of ancestral trait known as atavism.

↓

Example- Lion boy

Repetitive sequence in genome

Genome wide repeats

- Dispersed through out the genome

Examples:

SINE (Short Interspersed Nuclear Elements)

- ALU → Contain Alu Restriction site.
- Size 350 to 600 bp
- Almost 1 million copies present in genome

LINE (Long Interspersed Nuclear Elements) –

- L1 elements - Contain L1 endonuclease restriction site
- Size almost 6000 bp
- Almost 60,000 copies of line is present throughout the human genome

tandem repeats

present in constitutive heterochromatin region located on telomere and centromere.

- AT rich sequence
- Density is decreased than total genome

- Satellite DNA AT rich sequence - low density

Note human genome size is increased mainly due to **LINE** and SINE sequence i.e. a retro transposon.

Satellite DNA

mini satellite

- size 7 to 25 bp
- example TTAAGCT

micro satellite

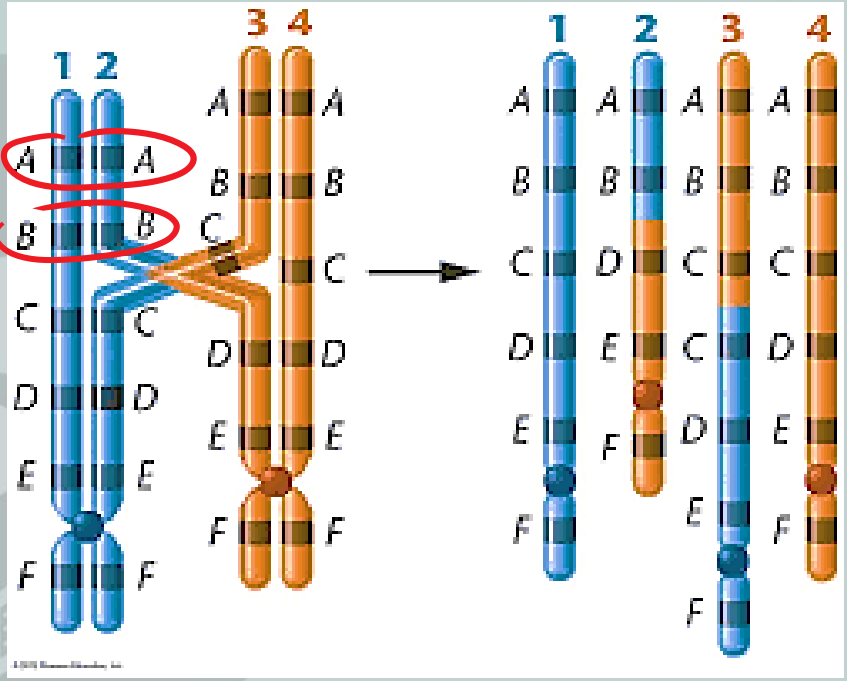
- size 1 to 6
- can be mononucleotide
- repeats or dinucleotide repeats

Satellite DNA is polymorphic in nature. – variation in nr in No.

Micro satellite is more polymorphic than mini satellite

Reason for polymorphisms-

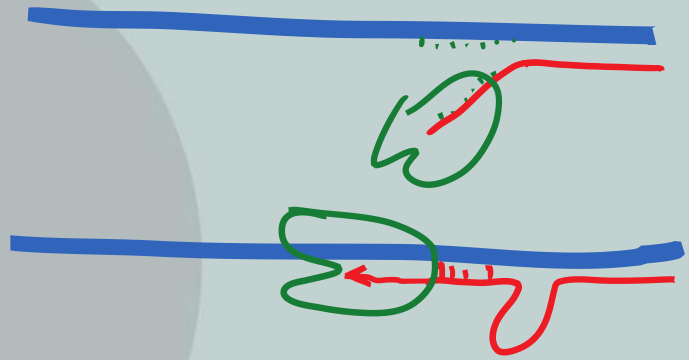
1. Unequal crossing over
2. Slippage error during replication



Slippage error occurs during replication

- – mostly occurs in telomeric region (random Repeats)
- - Generally it is not harmful

Exception- CAG repeats present in genic region (Lamin A).



- If CAG repeats increase more than 40 repeats, mutation occurs in Lamin A gene.
- It will lead to Hutchinson-Gilford Progeria

Micro satellite is because of polymorphism, used in DNA fingerprinting.

Example VNTR i.e. variable number of tandem repeats.

.

- Importance of polymorphism in genetic studies.

↓

- DNA finger printing
- marker (Genetic)

Facts about human genome

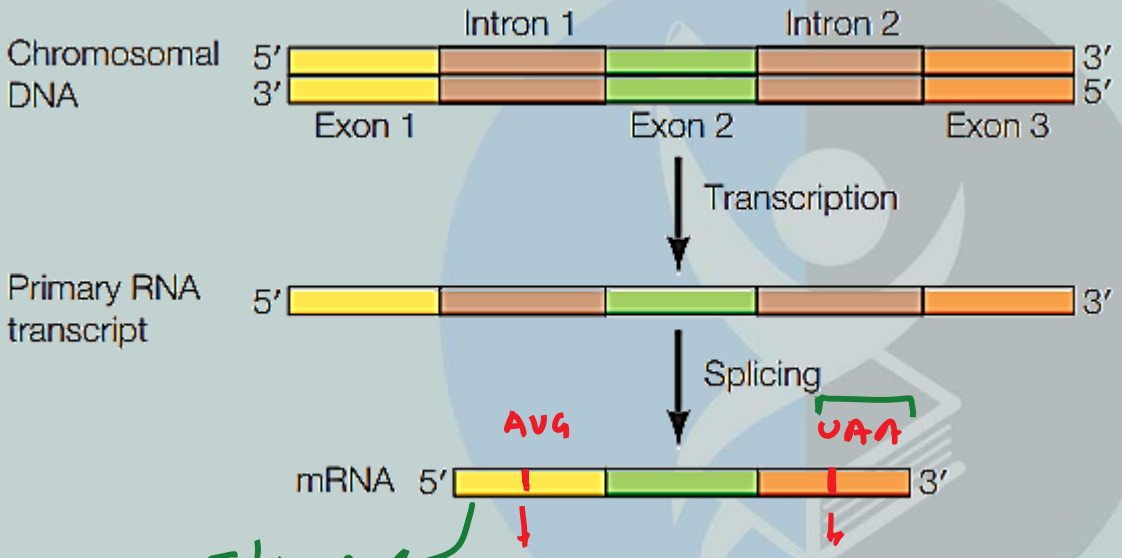
Human genome	
DNA length	3.2×10^9 nucleotide pairs*
Number of genes coding for proteins	Approximately 21,000
Largest gene coding for protein	2.4×10^6 nucleotide pairs
Mean size for protein-coding genes	27,000 nucleotide pairs
Smallest number of exons per gene (Histone)	1
Largest number of exons per gene	178
Mean number of exons per gene	10.4
Largest exon size	17,106 nucleotide pairs
Mean exon size	145 nucleotide pairs
Number of noncoding RNA genes	Approximately 9000**
Number of pseudogenes***	More than 20,000
Percentage of DNA sequence in exons (protein-coding sequences)	1.5%
Percentage of DNA in other highly conserved sequences****	3.5%
Percentage of DNA in high-copy-number repetitive elements	Approximately 50%

* The sequence of 2.85 billion nucleotides is known precisely (error rate of only about 1 in 100,000 nucleotides). The remaining DNA primarily consists of short sequences that are tandemly repeated many times over, with repeat numbers differing from one individual to the next. These highly repetitive blocks are hard to sequence accurately.
** This number is only a very rough estimate.
*** A pseudogene is a DNA sequence closely resembling that of a functional gene, but containing numerous mutations that prevent its proper expression or function. Most pseudogenes arise from the duplication of a functional gene followed by the accumulation of damaging mutations in one copy.
**** These conserved functional regions include DNA encoding 5′ and 3′ UTRs (untranslated regions of mRNA), DNA specifying structural and functional RNAs, and DNA with conserved protein-binding sites.

Largest gene is dystrophin but it is not code for largest protein.

Largest protein is titin.

Human gene is known as split gene -Exon is separated by intron

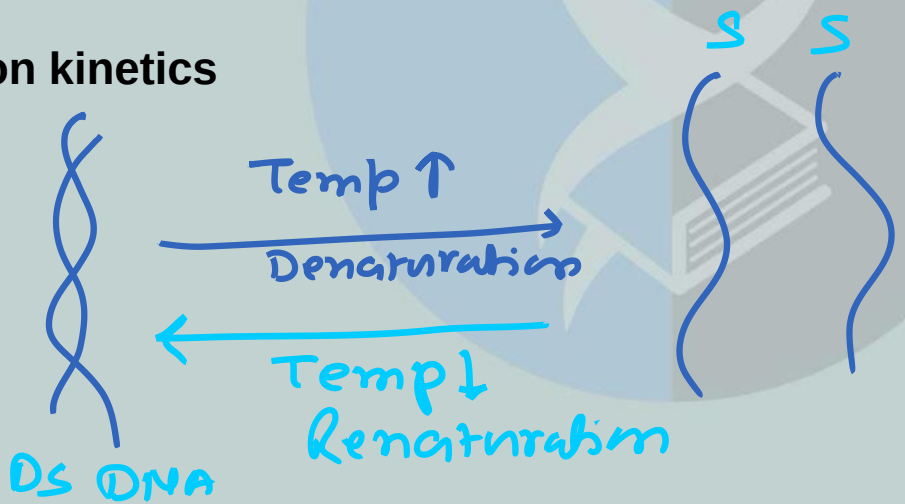


Understanding Genome Complexity

Genome complexity is determined by Cot value.

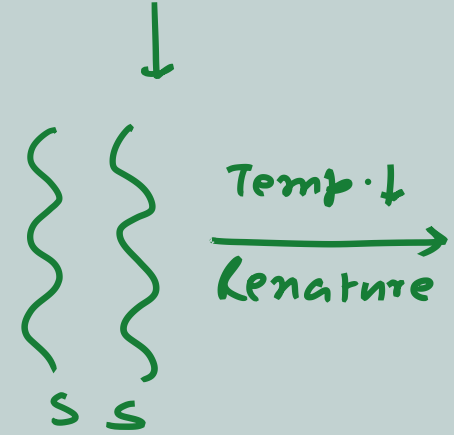
Renaturation kinetics

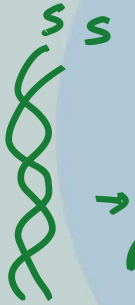
Denaturation kinetics



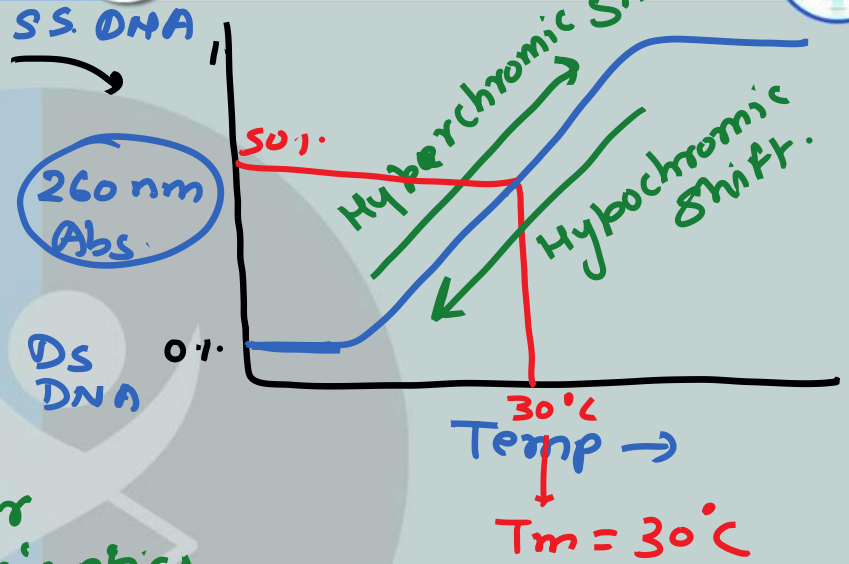
Tm = 50% of total DNA get Denatured

Renaturation kinetics-





→ 2nd order Reaction kinetics



Used to Determine –

- Repetitive DNA
- Genome complexity
- Genome size in prokaryotes.

Renaturation kinetics - can be determined by Cot1/2.

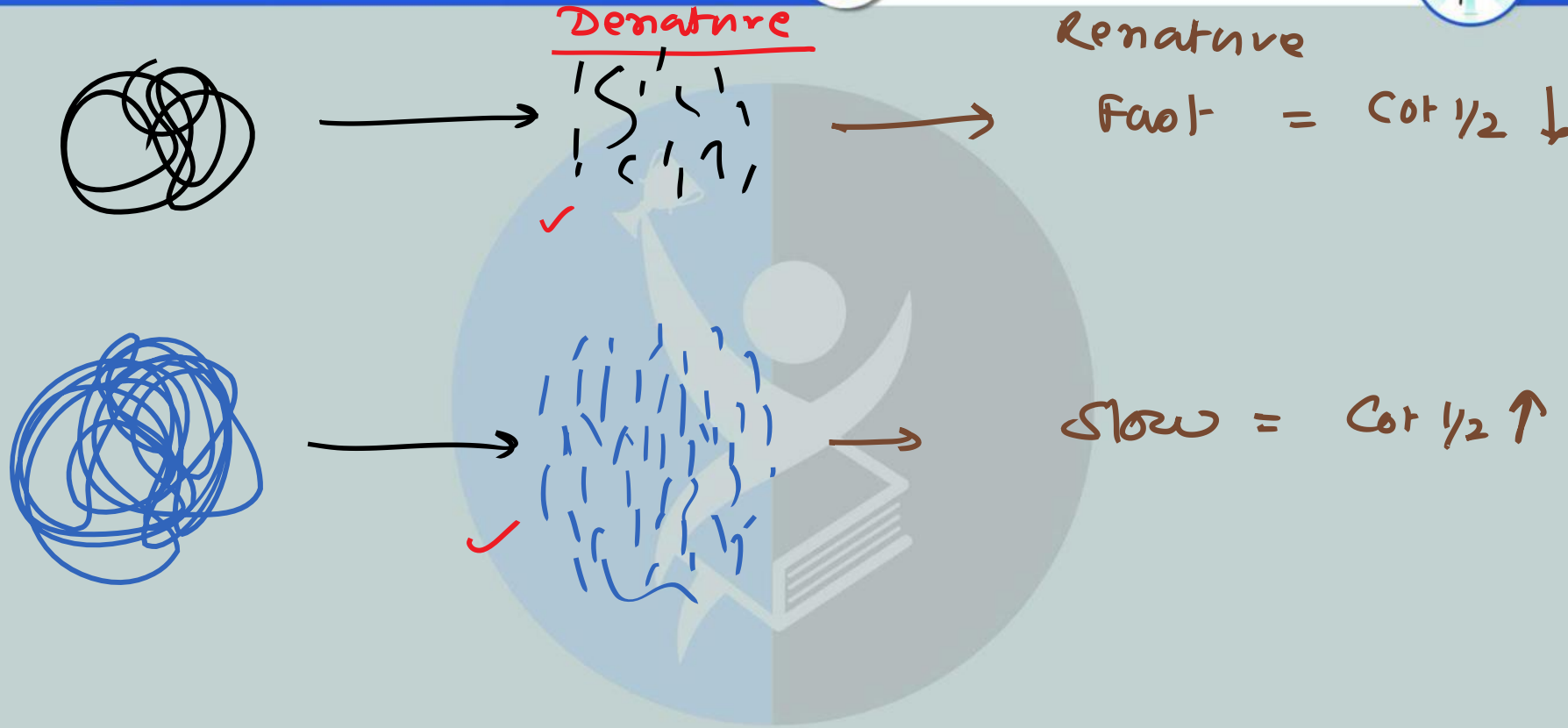
Cot1/2 = Time of incubation at which 50% of single-stranded DNA get converted into double-stranded DNA

Co = concentration of DNA.

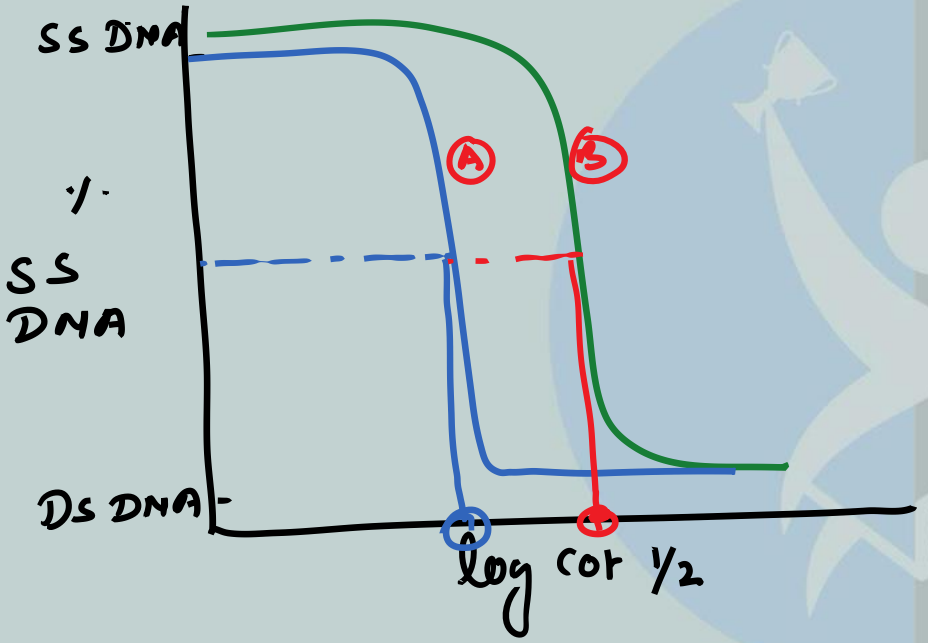
t = time of incubation (for renaturation).

Genome -

- Cot1/2 decrease = Renaturation increase.
- COt1/2 increase =Renaturation decrease


Denature
Renature
Fast = Cot 1/2 ↓
Slow = Cot 1/2 ↑

Prokaryotic organism



Organism B has larger genome size as comparison to organism A.

- Reason - B has high COT1/2 value.
- It will takes more time to renature.

Every prokaryotic organism has single COt value.

Eukaryotic organism

Highly repetitive sequence

High renaturation = low Cot1/2

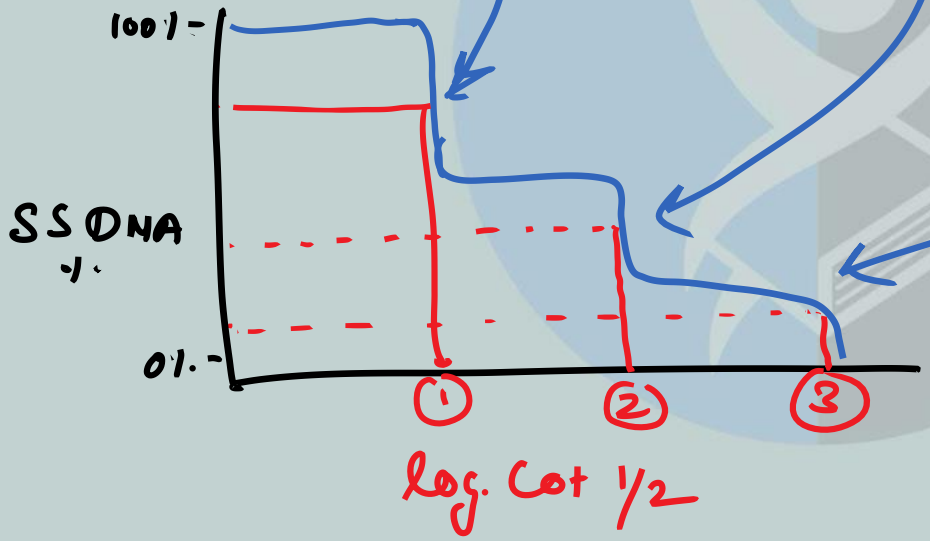
moderate repetitive sequence

moderate renaturation = moderate COT1/2

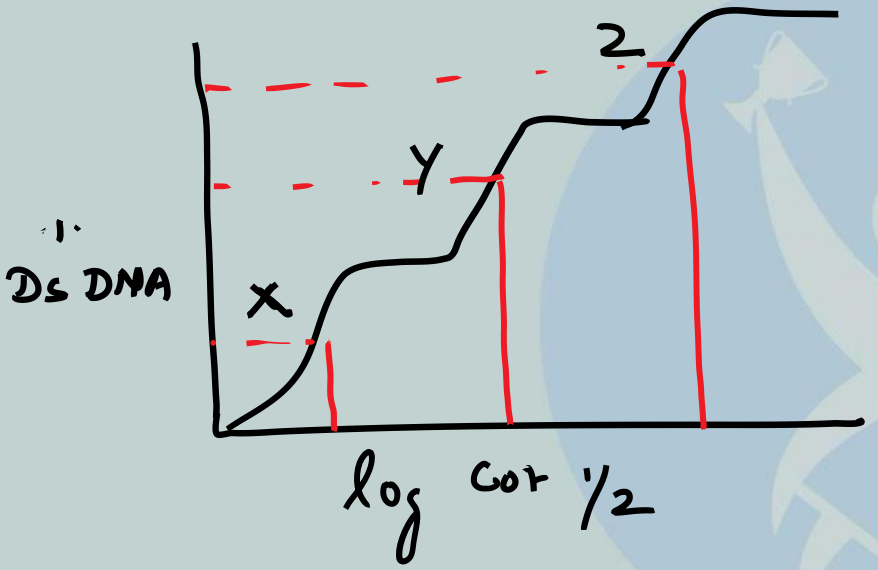
unique sequence

slow renaturation =

high COT1/2



Eukaryotic genome.



A = highly repetitive sequence.

B = moderate repetitive sequence

C = unique sequence

X = A

Y = B

Z = C

Application of C_0t value.

- Understanding of genome complexity
- genome size among prokaryote.
- Relative proportion of repetitive and unique sequence

www.ifasonline.com

thank you

Learn From India's Best Educators

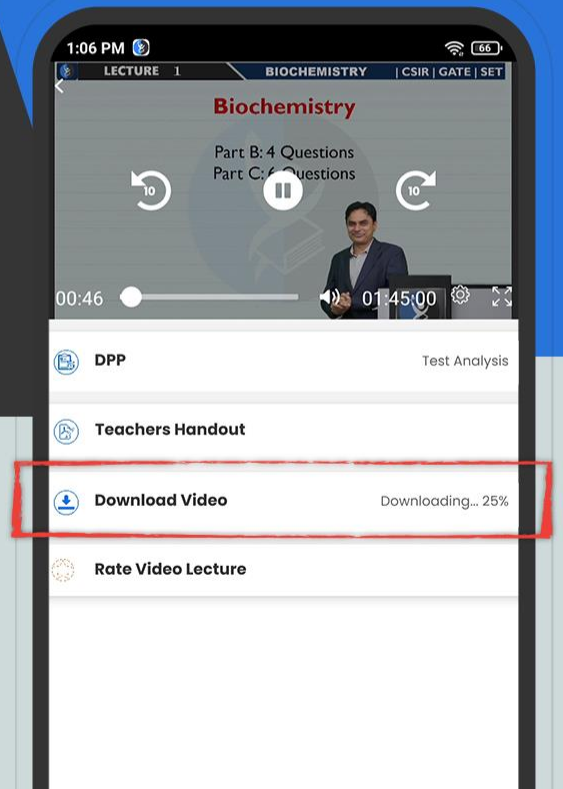
India's **No.1** EdTech Company for Graduates
& Post Graduates Examination.

Download **IFAS** App Now

ANDROID IOS WINDOWS

SAVE YOUR INTERNET DATA USAGE!

Minimize Internet Usage by Downloading Lectures from the IFAS App & Access them Offline



IfAS

Harshada Sharma

Earned: ₹25,500

"As is widely recognized, IFAS stands as India's leading institute for CSIR NET examination preparation. My Personal philosophy has always revolved around helping others and I've had the pleasure of introducing the IFAS Online Course to 12 of my classmates and juniors, which in turn allowed me to earn 25,500 rupees. You too can join in and refer your friends, allowing you to earn a 5% rewards based on the fees each of your referred friends Pay"

REFER & EARN

Step – 01 Invite Friends

Share your referral code or Link with your friends.

Step – 02 Earn Cashback

On Every Purchase made using your referral code, you earn 5% instant cashback in your bank account.

Step – 03 Refer More, Earn More

Keep referring as many friends as you can and keep **Earning** cash as much as you can

Friendship is Sharing, Caring, Giving & Helping Each Other...!